This article explores the distribution of to-infinitival complements (to-ICs) in Belfast English (BE) on the basis of Henry’s (1992) data in order to capture two generalizations about Standard English (SE) to-ICs: the complementizer for is incompatible with PRO, and complementizerless to-ICs are TPs. Importantly, these generalizations cannot be drawn from Henry’s CP analysis of BE to-ICs constructed within a representational approach. Thus, I advance a new analysis within the current minimalist framework (Chomsky (2000, 2001a, b)) on the basis of the idea that BE to-ICs have for which merges in T as a part of infinitival morphology and raises to C in the same way as V raises to v. I also propose BE-specific conditions on the selection of the categorial status of to-ICs. Further, I show two ways of capturing the nature of the adjacency condition within the minimalist framework.*

Keywords: to-infinitival complement, complementizer, PRO, null Case, infinitival morphology, adjacency, intervention effects

1. Introduction

Chomsky (1981) assumes that the complementizer for assigns Case to a to-infinitival subject, permitting a lexical subject and excluding PRO. Chomsky (1995, 2000, 2001a, b) assumes that PRO has null Case and
infinitival T (i.e. to) checks or values and deletes null Case. If we assume that for cannot assign null Case, then for and PRO are not compatible in to-infinitival clauses:

(1) a. *John tried [for PRO to study linguistics].
   b. *John preferred [for PRO to attend the meeting].
   c. *It is important [for PRO to arrive on time].

Boškovic (1997) provides a uniform IP analysis (=TP analysis within the current minimalist framework) for complementizerless to-infinitival complements (henceforth, COMP-less to-ICs) other than those of want-type verbs.1 Iwakura (2000) criticizes Boškovic’s analysis and advances an alternative analysis by proposing that all COMP-less to-ICs including those of want-type verbs are TPs, thereby providing a more uniform generalization.2 (2a) gives a control verb, (2b) an ECM verb, (2c, d) want-type verbs, (2e) a raising verb, and (2f) an adjective:

(2) a. John tried [TP PRO to study linguistics].
   b. I believe [TP John to be reliable].
   c. John wants [TP PRO to take a break].
   d. John wants [TP Mary to stay out].
   e. Johni appears [TP ti to fall for Mary].
   f. It is illegal [TP PRO to smoke here].

As pointed out by Boškovic (1997) and Iwakura (2000), such a TP analysis is favored over any CP analysis by Boškovic’s (1997: 25) Minimal Structure Principle (MSP), a modified version of Law’s (1991: 25):

1 With respect to want-type verbs, Boškovic assumes that their to-ICs have IP status when taking PRO subjects, whereas they have CP status when taking lexical subjects:

( i ) a. John wants [IP PRO to take a break].
   b. John wants [CP $\phi_{for}$ [IP Mary to stay out]].

Boškovic, following Watanabe (1993), assumes that the CP to-IC in (ib) is headed by a null complementizer for, which is generated under I, forming a for-to complex, and raises to C. Boškovic (1997: 19) states that Watanabe’s analysis follows a suggestion by Noam Chomsky. Similarly, Martin (2001) also argues for the null for in C as in (ib).

2 Iwakura (2000) assumes that Case is assigned by V or P to an adjacent NP (Chomsky (1981: 94)) and proposes that want-type verbs with lexical subjects in their to-ICs have the capacity to assign two values of Acc. Importantly, Iwakura claims that Case-assignment is divorced from Chomsky’s (2001a) agreement (cf. sec. 4.2).
282) Principle of Economy of Representation:3, 4

(3) \textit{The Minimal Structure Principle (MSP)}:

Provided that lexical requirements of relevant elements are satisfied, if two representations have the same lexical structure and serve the same function, then the representation that has fewer projections is to be chosen as the syntactic representation serving that function.

In this respect, the TP analysis of COMP-less to-ICs is more in the direction toward minimalism without the expense of a superfluous CP projection. The fact in (1) and Bošković’s and Iwakura’s TP analyses lead us to the following generalizations concerning to-ICs in Standard English (SE):5

(4) a. The complementizer for and PRO are incompatible.

b. COMP-less to-ICs are TPs.6

However, (4a, b) do not seem to hold for to-ICs in Belfast English (BE), since as reported by Henry (1992) the complementizer for must appear in all to-ICs regardless of whether they take a PRO subject or a lexical subject, as exemplified in (5a–f). It is also worth noting that Henry shows that their subjects, unlike those in SE, can in some cases precede for, as illustrated in (5b, d) (and Henry claims in (5a), (5c), and (5e) as well):

(5) a. I tried [for to get them].

b. I believe [them for to have done it].

c. I want [for to meet them].

d. I wanted [Jimmy for to come with me].

e. John seems [for to be better].

f. It was stupid [for them to do that].


To account for the distribution of (5a–f), Henry provides a CP analysis

---


4 By the same token, Doherty (1997) argues that COMP-less finite complements (i.e. \textit{that}-less clauses) are also IPs (=TPs) in terms of the principle of economy.

5 In this article, I follow Bošković’s and Iwakura’s TP analyses of COMP-less to-IC and Bošković’s MSP.

6 Here I do not take into consideration COMP-less to-ICs involving \textit{wh}-phrases.
under the Government-Binding (GB) Theory for all to-ICs in BE by having recourse to three main operations, for-cliticization to I (=T), for-deletion, and CP-deletion. Henry assumes that whether each of these operations is optional or obligatory and which of these operations should apply depend on the type of predicate that selects the to-IC. Be that as it may, what is important for our purposes is that Henry’s CP analysis cannot capture generalizations (4a, b) regarding SE.

In this paper, I attempt to retain (4a, b) even for BE to-ICs by claiming that BE to-ICs have for as a part of infinitival morphology in T and that it raises to C, just as V raises to v, when CP to-ICs are selected by matrix predicates. I then suggest a TP analysis of to-ICs with for stranded in T, whereas I suggest a CP analysis of to-ICs with for raised to C. The suggested TP analysis not only can maintain (4a, b) in BE to-ICs, but also is more in accordance with the spirit of minimalism in that it can dispense with a superfluous CP projection and the superfluous operations such as for-cliticization, for-deletion, and CP-deletion that Henry assumes. I also discuss how to capture the nature of the adjacency condition within the minimalist framework. Importantly, I reanalyze the distribution of BE to-ICs in terms of the current derivational approach, the minimalist approach (Chomsky (2000, 2001a, b)).

The paper is organized as follows: section 2 observes the wide distribution of to-ICs in BE on the basis of Henry’s data; section 3 introduces Henry’s CP analysis and points out its problems; section 4 offers an alternative analysis; and section 5 forms a conclusion.

2. To-Infinitival Complements in BE

Henry (1992) furnishes rich data regarding the wide distribution of to-ICs in BE. As mentioned above, they have two major characteristics: they have an obligatory use of for and allow their subjects to precede for, neither of which is permitted in SE. With this in mind, let us observe the distribution of to-ICs with each predicate type from Henry’s data.

To begin with, try-type verbs select the following patterns of to-ICs:7

According to Carroll (1983), in Ozark English, which also has an obligatory use of for in to-ICs, (6c) is completely acceptable.
(6) a. I tried [for to get them]. (=5a))
b. *I tried [him for to go home].
c. *I tried [for him to go home]. ((a)-(c) Henry (1992: 283))

Believe-type (ECM) verbs take the following to-IC patterns:
(7) a. *I believe [for to have done it].
b. I believe [them for to have done it]. (=5b))
c. *I believe [for them to have done it].

((a)-(c) Henry (1992: 285))

Want-type verbs have to-IC patterns like the following:
(8) a. I want [for to meet them]. (=5c))
b. I wanted [Jimmy for to come with me]. (=5d))
c. *I wanted [for Jimmy to come with me].

((a)-(c) Henry (1992: 279, 284))

Raising verbs select the following to-IC patterns:
(9) a. John seems [for to be better]. (=5e))
b. *It seems [for John to be better].

((a), (b) Henry (1992: 285))

Non-verbal predicates take the following patterns of to-ICs:
(10) a. It was stupid [for to do that].
b. *It was stupid [them for to do that].
c. It was stupid [for them to do that]. (=5f))

((a)-(c) Henry (1992: 279, 284))

The important point to notice here is that only to-ICs selected by non-verbal predicates compel their lexical subjects to follow for in the same manner as in SE, as evident from (10b, c), which contrast with (7b, c) and (8b, c). Interestingly, the same pattern obtains for want-type verbs only when some material intervenes between them and their to-ICs’ subjects (cf. (8b, c)):
(11) a. I want very much [for him to get accepted].
b. *I want very much [him for to get accepted].

((a), (b) Henry (1992: 284))

Further, Henry reports that the negative element not typically surfaces after the infinitival marker to in negative to-ICs:
(12) a. I would prefer [them for to not go].
b. *I would prefer [them for not to go].

((a), (b) Henry (1992: 285–286))

Only in to-ICs with for preceding their lexical subjects such as those in (10c) and (11a), may not occur before or after to (with the former being preferred):
(13) a. I would prefer very much [for them not to go].
   b. I would prefer very much [for them to not go].

((a), (b) Henry (1992: 286))

3. Henry’s (1992) Analysis

Henry (1992) analyzes BE to-ICs on the basis of the PRO Theorem proposed by Chomsky (1981: 191) within the framework of GB Theory; PRO is un governed. Henry argues that to-ICs in BE are all CPs at D-structure because they always take the complementizer for. Henry’s analysis relies on three main operations, for-cliticization to I, for-deletion, and CP-deletion, and assumes that whether each of them is optional or obligatory and which of them should apply depend on the type of the embedding predicate.

First, with respect to to-ICs of try-type verbs, Henry suggests the following:

(14) a. For-cliticization to I is optional.8
   b. For-deletion is obligatory when for is adjacent to the verb at S-structure.9 (Henry (1992: 298–299))

By applying (14a, b) to the D-structures of (6a–c), we derive S-structures like (15a–c):10

(15) a. I tried [CP [IP PRO [I for to] get them]]
   b. *I tried [CP [IP him [I for to] go home]]
   c. *I tried [CP for [IP him to go home]]

(15a) is acceptable with PRO ungoverned by the matrix verb because of a CP boundary; (15b) is ruled out with the infinitival subject not Case-assigned by the matrix verb owing to a CP boundary; and (15c) is also eliminated with for undeleted by (14b).11 Thus, it is predicted

---

8 Henry (1992: 290–291) assumes that after for-cliticization for does not leave a trace by claiming that for, which appears in any infinitive, has no semantic contribution. In this regard, Henry follows Chomsky’s (1991) assumption that items that do not enter into semantic interpretation at LF do not leave traces.

9 Henry’s for-deletion is based on a rule of for-deletion in PF proposed by Chomsky (1981).

10 I illustrate S-structures (15a–c) along the lines of Henry’s analysis. The same is true for (17a–c), (19a–c), (22a, b), and (24a–c).

11 As pointed out by an anonymous EL reviewer, if Case-marking is accomplished in the syntax, Henry’s analysis will incorrectly predict that (i) is acceptable.
that only (6a)/(15a) is grammatical.

Secondly, regarding to-ICs of ECM verbs, Henry proposes the following:

(16) a. For-cliticization to I is optional.
b. CP-deletion is obligatory.\(^{12}\) (Henry (1992: 298))

Applying (16a, b) to the D-structures of (7a–c), we have S-structures (17a–c):

(17) a. *I believe \([\text{IP}\ PRO \ I\ for\ to\ have\ done\ it]\)
b. I believe \([\text{IP}\ them\ I\ for\ to\ have\ done\ it]\)
c. *I believe \([\text{CP}\ for\ [\text{IP}\ them\ to\ have\ done\ it]\])

(17a) is excluded with PRO governed by the matrix ECM verb; (17b) is acceptable with the infinitival subject Case-assigned via ECM; and (17c) is ruled out with CP undeleted by (16b). Thus, it is predicted that only (7b)/(17b) is grammatical.

Thirdly, concerning to-ICs of want-type verbs, Henry advances the following:

(18) a. For-cliticization to I is optional.
b. CP-deletion is optional.
c. For-deletion is obligatory when for is adjacent to the verb at S-structure. (Henry (1992: 296–297))

Applying (18a–c) to the D-structures of (8a–c) derives S-structures (19a–c):

(19) a. I wanted \([\text{CP}\ \ [\text{IP}\ PRO \ I\ for\ to\ meet\ them]]\)
b. I wanted \([\text{IP}\ Jimmy\ I\ for\ to\ come\ with\ me]\)
c. *I wanted \([\text{CP}\ for\ [\text{IP}\ Jimmy\ to\ come\ with\ me]\])

(19a) is acceptable with PRO ungoverned by the matrix verb on account of a CP barrier; (19b) is also acceptable with the infinitival subject Case-assigned via ECM; and (19c) is eliminated with for undeleted by (18c). Thus, it is predicted that (8a, b)/(19a, b) are grammatical. Moreover, (18a–c) account for why for always precedes an in-

since for can assign Case to him before deleting in PF:

(i) I tried him to go home.

If this is correct, it will be an empirical shortcoming of Henry’s analysis, giving an additional reason for reanalyzing it.

\(^{12}\) Henry (1992: 298) states that CP-deletion must occur at some point in the derivation after for-cliticization has applied. However, I point out in fn. 13 that both operations must apply in the overt syntax.
finitival lexical subject only when some material intervenes between this type of verb and its to-IC, as observed in (11a, b), repeated here as (20a, b):

(20)  a.  I want very much [for him to get accepted].
    b.  *I want very much [him for to get accepted].

(20a) does not have to undergo for-deletion by (18c) since for is not adjacent to the matrix verb, so that (20a) is acceptable with the infinitival subject Case-assigned by for, while (20b) is excluded with the infinitival subject not Case-assigned by the matrix verb because of an intervening adverbial phrase (AdvP).

Fourthly, regarding to-ICs of raising verbs, Henry suggests the following:

(21)  a.  For-cliticization to I is optional.
    b.  CP-deletion is obligatory.  (Henry (1992: 298))

Applying (21a, b) to the D-structures of (9a, b) yields S-structures (22a, b):

(22)  a.  John seems [IP ti [I for to] be better]]
    b.  *it seems [CP for [IP John to be better]]

(22a) is acceptable with the legitimate raising of the surface subject by (21b), whereas (22b) is ruled out with CP undeleted by (21b). Thus, it is predicted that (9a)/(22a) is grammatical.

With regard to to-ICs of non-verbal predicates, Henry does not mention anything explicit except for an optional for-cliticization:

(23)  For-cliticization to I is optional.  (Henry (1992: 290))

In this way, D-structures of (10a–c) proceed to S-structures (24a–c):

(24)  a.  it was stupid [CP [IP PRO [I for to] do that]]
    b.  *it was stupid [CP [IP them [I for to] do that]]
    c.  it was stupid [CP for [IP them to do that]]

(24a) is acceptable with PRO ungoverned by the matrix predicate because of a CP barrier; (24b) is excluded with the infinitival subject not Case-assigned by anything; and (24c) is acceptable with the infinitival subject Case-assigned by for. Thus, it is predicted that (10a, c)/(24a, 13 In order for the surface subject (i.e. John) to overtly raise to the matrix subject position in (22a), for-cliticization in (21a) and CP-deletion in (21b) must apply in the overt syntax before the subject raising; otherwise, the raising cannot be achieved on account of a CP barrier (cf. fn. 12).
c) are grammatical. Henry's optional *for*-cliticization to I is motivated by the PRO Theorem: PRO must not be governed by *for*. Henry claims this operation has several advantages. First, it accounts for the fact about negative to-ICs; the negative element *not* appears after the infinitival marker *to*, as observed in (12a, b), reproduced here as (25a, b):

(25) a. I would prefer [them for to not go].
   b. *I would prefer [them for not to go].

(25b) is eliminated since *for*-cliticization to *to* forms a single unit, thus preventing *not* from splitting the sequence *for to* (Henry (1992: 293)).

Likewise, the optional operation accounts for the acceptability of (13a), duplicated here as (26), by preventing it from applying in this case to form a single unit *for to*, thereby permitting *not* to emerge between *for* and *to*:

(26) I would prefer very much [for them not to go].

The same holds for the contrast between (27a, b) (Henry (1992: 296)): (27a, b) involve *for*-cliticization to *to*, hence hindering any AdvP from intervening between them. This eliminates (27b) but not (27a):

(27) a. I want [the boys definitely for to be there].
   b. *I want [the boys for definitely to be there].

Further, evidence that Henry adduces to show that *for* of *for to* is a complementizer comes from the fact that two occurrences of *for* are impossible, as exemplified in (28a, b) (Henry (1992: 284, 286)):

(28) a. *I want very much [for him for to get accepted].
   b. *[For him for to pay the mortgage] would be just as expensive.

The deviance of (28a, b) motivates Henry to adopt *for*-cliticization from C to I.

In addition, *for*-cliticization explains the possibility of wh-subject-extraction from within to-ICs of *want*-type verbs, as shown in (29a), which is excluded in SE as a violation of the ECP since *for* in C blocks antecedent government of the wh-trace in Spec-I by the intermediate wh-trace in Spec-C (Henry (1992: 296–297)):

(29) a. Who do you want [for to help you]?
   b. whoi do you want [CP t; [IP t; [I for to] help you]]

Adopting *for*-cliticization to I, (29a) has the S-structure (29b) in which *for* no longer blocks the antecedent government, so that (29a) is acceptable.

However, Henry's CP analysis faces several problems. First, it is
difficult to explain why the three operations Henry invokes, for-cliticization, CP-deletion, and for-deletion, are sometimes optional and sometimes obligatory, depending on matrix predicate types. This implies that these operations have a theoretically undesirable “look-ahead” property; that is, in order to determine whether they are optional or obligatory and which of them should apply, we need to know PF and LF outputs in advance. Hence, this raises serious issues of globality.\footnote{Bosković (2000: 80) points out that a similar problem arises with respect to Chomsky’s (1995: 294) economy principle in (i), since in order to determine the effect of (i) we need to know PF and LF outputs.}

Further, as remarked in section 1, if we adopt Henry’s CP analysis of BE to-ICs, we lose the uniform generalizations (4a, b) from SE to-ICs; that the complementizer for and PRO are incompatible and COMP-less to-ICs are TPs. To maintain (4a, b) in BE as well as in SE, I will present an alternative TP analysis of BE to-ICs in the next section, in which for of the sequence for to is not a complementizer. The TP analysis will resonate with several strands of current research moving toward economy driven approaches to syntax (Bošković (1997), Doherty (1997), Iwakura (2000), etc.); the TP analysis will be preferred to any CP analysis by economy principles like Bošković’s MSP in that it can do away with a superfluous CP projection, consistent with the spirit of minimalism. Likewise, if it is possible to do without Henry’s three operations, this is a desirable outcome. Hence, it is necessary to account for BE to-ICs in terms of the current minimalist approach, not a representational approach like Henry’s, without resort to D/S-structures, the PRO Theorem, the concept of government, etc.

4. An Alternative Analysis

In this section, I advance an alternative analysis to capture the parallelism between to-ICs in BE and those in SE (i.e. generalizations (4a, b)) in conformity with Chomsky’s (2000, 2001a, b) agreement theory. I argue that BE to-ICs have for which merges in T as a part of infinitival morphology, not in C as a complementizer, and that it affixes onto C in the same manner as V affixes onto v when CP to-ICs

\footnote{Bošković (2000: 80) points out that a similar problem arises with respect to Chomsky’s (1995: 294) economy principle in (i), since in order to determine the effect of (i) we need to know PF and LF outputs.}

(i) a enters the numeration only if it has an effect on output.
are selected by matrix predicates (see sec. 4.1). As pointed out in fn. 1 and Iwakura (1997: 162, fn. 7), Chomsky makes the same suggestion regarding SE want-type verbs’ to-ICs with lexical subjects. Chomsky (2001a: 9–12) assumes that what is common to CP and vP is that substantive categories are selected by functional categories, i.e. V by v and T by C, and that the C-T relation is therefore analogous to the v-V relation. This makes it plausible to assume that if V raises to v, for raises to C. Thus, to-ICs with for stranded in T are TPs headed by the complex for to, as illustrated in (30a), while those with for raised to C are CPs, as exemplified in (30b):

(30)  
\begin{itemize}
  \item a. \[\text{TP SUBJ} \rightarrow [T \rightarrow \text{for} \rightarrow] \rightarrow \ldots\]
  \item b. \[\text{CP} \rightarrow [\text{TP SUBJ} \rightarrow \text{to} \rightarrow] \rightarrow \ldots\]
\end{itemize}

It should be noted that for of for to in (30a) has no Case agreeing ability as long as it is only a part of infinitival morphology from the outset, while for in (30b) displays Case agreeing ability once independent in C. I assume that for-affixation to C applies only when matrix predicates select CP to-ICs in the same fashion as V-to-v affixation applies in vPs. Notice that (30a) corresponds to COMP-less TP to-ICs in SE (cf. (2a–f)). On the basis of (30a, b), I propose the following BE-specific conditions on the selection of the categorial status of to-ICs:

(31)  
\begin{itemize}
  \item a. Verbal predicates other than want-type verbs select TP to-ICs (i.e. (30a)).
  \item b. Want-type verbs select TP to-ICs (i.e. (30a)) when there is no element that is closer to v/V than the to-ICs.
\end{itemize}

I assume that these conditions are satisfied at matrix vP/CP phases before Spell-Out (cf. Chomsky (2001b: 21–22)) and that (31b) is indirectly derived from Chomsky’s (2000, 2001a, b) intervention effects or Iwakura’s (2000) Case-assignment analysis as discussed in sec. 4.2 and 4.3. The important point is that (31a, b) imply that in all other cases either TP or CP to-ICs (i.e. either (30a) or (30b)) are selected as a free option, but whether or not each of their derivations converge is determined in a principled way under Chomsky’s agreement theory.

(31a) is motivated by the fact that try-type verbs, ECM verbs, and raising verbs in BE never take CP to-ICs headed by for, as shown in (6c), (7c), and (9b), reproduced here as (32a–c) (cf. (8c)):

(32)  
\begin{itemize}
  \item a. *I tried [for him to go home].
  \item b. *I believe [for them to have done it].
  \item c. *It seems [for John to be better].
\end{itemize}

It is important to notice that these verbs never take CP to-ICs headed
by *for even if some element intervenes between them and their to-ICs, as illustrated in (33a–c):15

(33)  
   a. *I tried very hard [for him to go home].
   b. *I believe strongly [for them to have done it].
   c. *It seems evidently [for John to be better].

Note that the SE counterparts of (32a–c) and (33a–c) are not grammatical either. The deviance of (32a–c) and (33a–c) gives support for (31a).

(31b) is motivated by the fact that want-type verbs never take CP to-ICs headed by *for when they are the closest to their to-ICs, as evident from the contrast between (8c) and (11a), repeated here as (34a, b):

(34)  
   a. *I wanted [for Jimmy to come with me].
   b. I want very much [for him to get accepted].

(Cf. SE: I wanted [for Jimmy to come with me].)

It is worth noting that the deviance of (34a) shows that (34a) cannot be analyzed in the same way as its SE equivalent under Bošković’s (1997) and Martin’s (2001) CP analyses (cf. fn. 1).

If the proposed analysis is empirically correct, then it has the advantage over Henry’s analysis of accounting in a simpler way for the distribution and the categorial status of BE to-ICs. Further, given structure (30a), the analysis will enable us to retain generalizations (4a, b) from SE, repeated here as (35a, b), in BE as well:

(35)  
   a. The complementizer *for and PRO are incompatible.
   b. COMP-less to-ICs are TPs.

4.1. *For as a Part of Infinitival Morphology

We have thus far observed that BE to-infinitival (PRO and lexical) subjects can precede *for, as in (6a), (7b), (8a, b), (9a), (10a), etc., whereas the lexical subjects follow *for only when to-ICs are selected by non-verbal predicates, as in (10c), and when they are not adjacent to matrix verbs because of some intervening element, as in (11a), (13a, b), etc. In the former case, i.e. [SUBJ for to ...], Henry assumes the optional overt *for-cliticization from C to I to acquire the sequence *for to and to avoid a violation of the PRO Theorem; PRO must not be governed by *for. However, I claim that *for in this case is not a com-

---

15 I am grateful to an anonymous EL reviewer for drawing my attention to examples like (33a–c).
plementizer but only a part of infinitival morphology from the beginning, contrary to Henry’s analysis. In the latter case, i.e. [for SUBJ to ...], I assume that for affixes onto C, as mentioned above. We can now dispense with Henry’s optional for-lowering (as well as for-deletion and CP-deletion).

My claim is based on historical facts about to-infinitives; as pointed out by Mustanoja (1960), Visser (1963–73), Nakao (1972), etc., for was introduced in the late 11th century as an intensifier of to to reinforce the directive meaning of to, and later for to and to became free variants. This makes it plausible to consider that for to and to are placed in the same position. According to the Oxford English Dictionary (OED), it is also suggested that for to is simply a lexical variant of to in some varieties. Further, similar to my claim, Lightfoot (1979: 186–188, 195–196) also argues that for of for to is part of infinitival morphology, which is evident from the fact that a to-infinitival subject could have nominative Case and precede for, as shown in (36b):

(36) a. [For to go] is necessary. (1205–1590)
   b. [I for to go] is necessary. (14th c.–1607)
   c. [For us to go] is necessary. (1567)

This shows that for in (36a, b) may be different from for in (36c) in its syntactic status, a possibility which is greatly strengthened by the fact that the former was made use of much earlier than the latter, as pointed out by Lightfoot (1979: 195). Thus, it is plausible to consider (30b) to be derived from (30a).

More convincing evidence of for in (30a) serving as part of infinitival morphology comes from the historical fact that for and to are often employed as one word, as illustrated in (37a–d):

(37) a. Forte don him understonden.
    (a1175 Cott. Hom. 211 (OED))
   b. it is no shame forto swinken
    (c1300 Hauelok 799, in Lightfoot (1979: 201))

Mustanoja (1960: 514) suggests that for to became a mere sign of the infinitive, equivalent to to, in the course of the 13th century, when it became quite common. Mustanoja states that the appearance of for to has been ascribed to French and also to Scandinavian influence, though the matter is by no means clear.

Lightfoot (1979: 187) points out that “although several for clause-types are represented ..., the for NP to V construction appears consistently about 200 years later than the corresponding for to V forms.”
c. Seint swythin ... swipe ʒung bigan Forto seruie ihesu crist. (c1305 St. Swithin 14 in E.E.P (1862) 43 (OED))
d. ʒif eni mon bit forti iseon ou
   (Ancr. 41, in Mustanoja (1960: 515))

This demonstrates that for and to merge as one unit in the same place, thus justifying (30a).

In addition, my claim that for in (30a) is a part of infinitival morphology from the outset is supported by the fact that some dialects such as Jamaican Creole have an infinitival marker for merging in T instead of to, as illustrated in (38a, b), adapted from Bailey (1966):

(38) a. Mi waan yu fi sel i.
    me want you for sell it
    ‘I want you to sell it.’

b. Im waan fi haid dem.
    him want for hide them
    ‘He wants to hide them.’

From these facts, we are in a position to assume (30b) to be derived from (30a) by means of for-affixation onto C.

The proposed analysis is favored over Henry’s analysis in excluding for-cliticization (for-deletion and CP-deletion as well) and avoiding the question of why the operation is optional, thus eliminating the theoretically undesirable “look-ahead” property. Further, as is well known, the PRO Theorem Henry considers to be one of the reasons for for-cliticization is no longer adopted within the current minimalist framework in which the concept of government as well as the representational levels of D/S-structure are eliminated and PRO has null Case to be valued and deleted by infinitival T (cf. sec. 4.2). It is also questioned whether such an overt lowering rule as for-cliticization can be regarded as legitimate within the minimalist framework. Accordingly, it is essential to provide a simpler alternative analysis of BE to-ICs in minimalist terms.

4.2. Chomsky’s (2000, 2001a, b) Agreement Theory and Adjacency

Chomsky (1995) suggests that infinitival T (i.e. to) checks the null Case of PRO. Chomsky (2000: 122) takes structural Case to be a reflex of an uninterpretable $\phi$-set and Chomsky (2001a: 6) further assumes that structural Case is not a feature of the probes (T, v) but deletes under matching of $\phi$-features, though it itself is not matched. According to Chomsky (2001a: 6), the operation Agree is subject to
the following principles:

(39) a. Goal as well as probe must be active for Agree to apply.
b. \( \alpha \) must have a complete set of \( \phi \)-features to delete uninterpretable features of the paired matching element \( \beta \).

(39a, b) are based on Chomsky's (2001a: 4) assumption that uninterpretable features of the probe and goal render their relevant subparts active. Given (39a, b), it follows as pointed out by Chomsky (2001a: 7-8) that a \( \phi \)-complete T or \( \nu \) values and deletes the uninterpretable structural Case feature of N and that the \( \phi \)-set of N, which is always \( \phi \)-complete, values and deletes the uninterpretable \( \phi \)-features of T or \( \nu \).

Note that in this case both T/\( \nu \) and N, having uninterpretable features, are active and that the structural Case feature of N is not matched but deletes under matching of \( \phi \)-features with T/\( \nu \), as mentioned above.

Let us now consider how to value and delete the null Case of PRO under matching of \( \phi \)-features with infinitival T by strict adherence to (39a, b). Chomsky (1995: 119) assumes that PRO, like other NPs, has \( \phi \)-features as well as a Case feature,\(^{18}\) and Chomsky (2001a: 9) assumes that raising T and ECM T are defective, whereas control T is \( \phi \)-complete. This leads us to assume that control T is the only probe capable of valuing and deleting the null Case of PRO under matching of \( \phi \)-features. Following this assumption, we can account for the following contrast in SE by saying that the \( \phi \)-complete infinitival T in (40a) can value and delete the null Case of PRO, leading the derivation to converge, while the defective infinitival T in (40b) cannot, thus inducing the derivation to crash:

(40) a. John tried [PRO to study linguistics].
b. *I believe [PRO to be a liar].

However, Chomsky's Case-agreement under matching of \( \phi \)-features seems to be problematic with regard to SE examples (and their BE counterparts) like the following, which have been excluded as a violation of the adjacency condition on Case-assignment within the GB

---

\(^{18}\) Terzi (1992) claims that PRO has \( \phi \)-features by showing that infinitival T in Balkan subjunctives allows PRO, although they exhibit the same verbal morphology as indicatives (Bošković (1997: 10, 177, fn. 4)):

( i ) O Yiannis theli PRO na fai.
Yiannis wants SUBJ.PART eats
‘Yiannis wants to eat.’
framework (Chomsky (1981: 94)):

(41) a. *I believe sincerely him to be honest.
    b. *It is important for tomorrow her to go to Kochi.
    c. *Mary likes very much me.

Under Chomsky's Case-agreement analysis, in (41a–c), the probe v/for/v values and deletes the Case feature of the goal him/her/me under matching of φ-features, leading the derivation to converge. The important point here is, as pointed out by Oshima (2001: 86) (and in personal communication), that the intervening AdvP, which has neither Case nor φ-features, does not block Agree between the probe and the goal since Agree is defined in terms of active features and complete φ-features, not in terms of adjacency. In other words, AdvPs do not yield intervention effects. If so, we incorrectly predict that examples like (41a–c) are acceptable.

However, whether AdvPs have φ-features appears to be debatable. Carnie (2002: 36) uses the syntactic category A for both adjectives and adverbs by pointing out that “in much work on syntactic theory, there is no significant distinction between adjectives and adverbs ... there are other distributional criteria that suggest they might be the same category. They both can be modified by the word very, and they both have the same basic function in the grammar—to attribute properties to the items they modify. The issue is still up for debate.” If we maintain Carnie's position, we can assume that adverbs have the categorial feature [+N, +V] in the same way as adjectives. This enables us to suppose that adverbs may have φ-features and thus serve as candidates to create intervention effects. Hence, I tentatively suggest the following in order to capture the essence of the adjacency condition within the minimalist framework:

(42) In English adverbs may have φ-features, thus yielding intervention effects on Agree between the probe (v, P, C(for)) and the goal NP/DP.

---

19 This is also pointed out by an anonymous EL reviewer.
20 I am grateful to Hiroshi Hasegawa and Ken-ichi Takami for pointing out this line of reasoning.
21 If whether adverbs contain the categorial feature [+N] or [−N] is still not clear, as pointed out by Carnie (2002: 36), then it is possible to assume that they contain [+N] in certain cases and [−N] in other cases, or that they have φ-features.
Following (42), we can eliminate (41a–c) by saying that the intervening AdvP prevents Agree between the probe \(v/for/v\) and the goal \(him/her/me\), yielding an intervention effect and thus causing the derivation to crash. Accordingly, we can correctly predict that (41a–c) are deviant. If there is no assumption like (42), the deviance of examples like (41a–c) will remain unaccountable under Chomsky’s agreement theory. In this respect, I assume that (42) is tenable.

Alternatively, Iwakura (2000: 219) claims that “Chomsky’s Case-assignment should be restricted to cases involving a probe and a goal which exhibit morphological agreement between them. In other cases showing no such agreement, there is no justifiable reason to suppose that Case … deletes under matching of \(\phi\)-features.” His claim is based on the fact that in (43a–c) the probe \(v/of/for\) can simply value and delete the Case feature of the goal him without involving matching of \(\phi\)-features because there is no morphological agreement between them:

(43)  
\begin{enumerate}  
\item They like him.  
\item They are fond of him.  
\item It is good for him to work harder. (Iwakura (2000: 219))  
\end{enumerate}

Hence, he suggests that “Case-assignment is divorced from agreement.” To support this view, he further points out that if Chomsky’s Case-agreement applies to (43b, c), we have to assume the preposition \(of\) and the complementizer \(for\) to have a complete set of \(\phi\)-features; otherwise, we cannot account for Case-agreement phenomena between Ps/C and Ns.\(^{22}\) From this argument, he suggests that “a lexical head with a Case-assigning property can value and delete the Case feature of an element without involving agreement of \(\phi\)-features.” To account for how a probe can find a matching goal under this assumption, he makes the point that “lexical Case-assigners are subject to the adjacency condition on Case-assignment”: a probe has a goal adjacent to it. With

in certain cases and do not in other cases. If this is correct, we can account for why they do not produce intervention effects on Agree between the probe \(T\) and the goal NP/DP in examples such as (i). This is because in this case they have \([-N]\) and no \(\phi\)-features, thus yielding no intervention effect:

\((i)\) There is/are AdvP NP …

I leave this issue for future research.

\(^{22}\) Note that Martin (2001) and Matsubara (2000) assume that the complementizer \(for\) and prepositions are \(\phi\)-complete.
respect to the locus of Case-assignment/EPP, he adopts \textit{Locus}_{TV} (which indicates that the locus is T, V). Consequently, he proposes (44) (Iwakura (2000: 226)):\footnote{Iwakura (2000) assumes that the expletive \textit{there} has a Case feature. Thus, he claims that the Case-bearing \textit{there} assumption is consistent with the adjacency condition, as shown in (i):

(i) We expect [there to be someone in the room]. (Chomsky (2001a: 16))
In (i) the probe \textit{expect} values and deletes the Case feature of the goal \textit{there} in accord with the adjacency condition.}

(44) When a lexical Case-assigner is adjacent to a nominal phrase (NP/DP), it can value and delete the Case feature of the nominal phrase in accordance with the adjacency condition. Applying (44) to (41a–c), we can rule out (41a–c) by saying that the probe \textit{believes/for/likes} cannot value and delete the Case feature of the goal \textit{him/her/me} since they are not adjacent to each other. Thus, we can correctly predict that examples like (41a–c) are deviant.

Now we have two ways of excluding examples like (41a–c): one is by intervention effects under Chomsky’s agreement theory and the other is by the adjacency condition under Iwakura’s Case-assignment analysis. I assume that both are applicable to BE examples corresponding to those like (41a–c) in the same manner. The choice between the two approaches does not affect the conclusion of this article, but I discuss theoretical consequences of the proposed analysis in the following sections along the line of the former approach.

4.3. Theoretical Consequences

I have suggested that BE \textit{to-ICs} have \textit{for} as a part of infinitival morphology in T and that \textit{for} undergoes affixation onto C in (30a). On the basis of this, I have also proposed that most clause embedding verbs select TP \textit{to-ICs} and that \textit{want-type} verbs select TP \textit{to-ICs} when there is no element that is closer to \textit{v/V} than the \textit{to-ICs}, as in (31a, b). Here I demonstrate that the proposed analysis can account in a simpler manner than Henry’s for the distribution of BE \textit{to-ICs} within the current minimalist framework. This enables us to maintain generalizations (4a, b)\textsuperscript{(=35a, b)} in BE as well as in SE; the complementizer \textit{for} and PRO are incompatible and COMP-less \textit{to-ICs} are TPs.
To start with, given the suggested analysis, the to-ICs of try-type verbs in (6a–c) have the following structures in the course of derivation:

\[(45)\]

\begin{enumerate}[a.]
\item I tried \([TP \ \text{PRO} \ [T \ \text{for} \ \text{to}] \ \text{get} \ \text{them}]\)
\item \*I tried \([TP \ \text{him} \ [T \ \text{for} \ \text{to}] \ \text{go} \ \text{home}]\)
\item \*I tried \([CP \ \text{for} \ [TP \ \text{him} \ to \ \text{go} \ \text{home}]\]
\end{enumerate}

(45a, b) select a (30a)-type of TP to-IC headed by the complex infinitival morphology for to, while (45c) selects a (30b)-type of CP to-IC in which for has affixed onto C. (45c) is ruled out as a violation of (31a) because (31a) requires that the verb take a TP to-IC. The same account holds for (33a), repeated here as (46):

\[(46)\] *I tried very hard \([CP \ \text{for} \ \text{him} \ to \ \text{go} \ \text{home}]\].

In (45b, c) and (46) \(\phi\)-complete control T, which is the only probe to value and delete the null Case of PRO, cannot value and delete the Case feature of him under matching of \(\phi\)-features, which causes the derivation to crash with the Case agreeing capacity of control T left intact.\(^{24}\) In (45a), on the other hand, control T with a complete \(\phi\)-set values and deletes the null Case of PRO, and the \(\phi\)-set of PRO values and deletes the \(\phi\)-features of control T. Hence, the derivation converges, yielding the acceptable example (6a). Therefore, the proposed analysis correctly predicts that only (6a) is acceptable. What is more important here is that the present analysis has advantages over Henry’s in that it can dispense with (14a, b) as well as the D/S-structure representations (15a–c) and in that the question of why (14a, b) are optional or obligatory is eliminated.

Secondly, according to the analysis here, the to-ICs of ECM verbs in (7a–c) have the following structures in the course of derivation:

\[(47)\]

\begin{enumerate}[a.]
\item \*I believe \([TP \ \text{PRO} \ [T \ \text{for} \ \text{to}] \ \text{have} \ \text{done} \ \text{it}]\]
\end{enumerate}

\(^{24}\) Alternatively, the derivation of (45b, c) and (46) will be canceled because of feature mismatch between control T and him with regard to Case (Chomsky (1995: 309)).

Bošković (1997: 16) excludes the SE counterpart of (45b) by saying that him must move to the matrix Spec-Agr\(_P\) for Case checking and this movement of him from a Case-checking position is ruled out by the Last Resort Condition. However, Bošković’s claim is not in accordance with the current minimalist framework where Agr(P) and English Object Shift for Case checking are abandoned (Chomsky (2000, 2001a, b)).
b. I believe [TP them [T for to] have done it]
c. *I believe [CP for [TP them to have done it]]

(47a, b) select (30a)-type of TP to-ICs headed by for to, whereas (47c) selects a (30b)-type of CP to-IC where for has undergone affixation onto C. (47c) is eliminated as a violation of (31a) because (31a) requires that the verb select a TP to-IC. Similar remarks apply to (33b), reproduced here as (48):

(48) *I believe strongly [CP for them to have done it].

(47a) is also excluded since defective ECM T is not capable of valuing and deleting any Case, thus inducing the derivation to crash with the null Case of PRO unvalued and undeleted. Notice here that the matrix v, although \( \phi \)-complete, cannot value and delete the null Case of PRO since it plays a role of valuing and deleting accusative (Acc) (Chomsky (2001a: 6)), thus leading the derivation to the same result.\(^{25}\)

In (47b) the Case feature of them is valued and deleted by v via ECM under matching of \( \phi \)-features, and the \( \phi \)-set of v is valued and deleted by the \( \phi \)-features of them, thus leading the derivation to converge. Therefore, the suggested analysis rightly predicts that only (7b) is acceptable. More importantly, the analysis here is preferred over Henry's in that it can eliminate (16a, b) as well as the D/S-structures (17a-c) and in that it does not raise the question of how to explain the optionality or obligatoriness of (16a, b).

Thirdly, following the analysis advanced, the to-ICs of want-type verbs in (8a-c) have the following structures in the course of derivation:

(49) a. I want [TP PRO [T for to] meet them]
   b. I wanted [TP Jimmy [T for to] come with me]
   c. *I wanted [CP for [TP Jimmy to come with me]]

(49c) is excluded as a violation of (31b) since the verb selects a (30b)-type of CP to-IC with for-affixation onto C, in spite of the fact that there is no element that is closer to v/V than the to-IC (cf. (34b)). What is worth noticing in (49a, b) is that want-type verbs share some properties with both try-type and ECM verbs: they allow both PRO

\(^{25}\) As pointed out in fn. 24, another possibility is that the derivation of (47a) will be canceled because of feature mismatch between the ECM v and PRO concerning Case, that is, accusative vs. null (Chomsky (1995: 309)).
and lexical subjects in their TP to-ICs. This implies that in the former
case want-type verbs select a $\phi$-complete control T, whereas in the latter
case they select a defective ECM T. Given this, we correctly pre-
dict that (49a) is as acceptable as (45a) and that (49b) is as acceptable
as (47b); in (49a) control T values and deletes the null Case of PRO
under matching of $\phi$-features, whereas in (49b) the matrix v values and
deletes the Case feature of Jimmy via ECM under matching of $\phi$-fea-
tures. Hence, their derivations result in convergence, yielding the
grammatical examples (8a, b). Here also, the present analysis suc-
ceeds in excluding (18a–c) as well as the D/S-structures (19a–c) and in
getting rid of the question of why (18a–c) are sometimes optional and
sometimes obligatory.

Likewise, the proposed analysis can also account for the fact that the
to-ICs of want-type verbs force their lexical subjects to follow for when
some material intervenes between the verbs and their to-ICs, as in
(11a, b), duplicated here as (50a, b):

(50) a. I want very much [for him to get accepted].
    b. *I want very much [him for to get accepted].

In (50a) the verb selects a (30b)-type of CP to-IC with $\text{for}$-affixation
onto C, while in (50b) the verb selects a (30a)-type of TP to-IC headed
by $\text{for}$. In compliance with (31b), we expect both cases to be allow-
able since when there is some element that is closer to v/V than its to-
IC, either CP or TP to-ICs are to be selected as a free option, as noted
above. However, the contrast between (50a) and (50b) can readily be
explained by (42) under Chomsky’s agreement theory (or by (44) under
Iwakura’s Case-assignment analysis), as argued in sec. 4.2. In (50a)
Agree holds between $\text{for}$ and him, valuing and deleting the $\phi$-features
of $\text{for}$ and the Case feature of him, thus leading the derivation to con-
verge (cf. fn. 22). On the other hand, according to (42), the interven-
ing AdvP in (50b), which has $\phi$-features, yields an intervention effect,
thereby blocking Agree between v and him. This causes the deriva-
tion to crash with the $\phi$-features of v and the Case feature of him un-
valued and undeleted. Consequently, we rightly predict that (50a) is
grammatical but (50b) is not. The same result obtains from (44).

Fourthly, given the suggested analysis, the to-ICs of raising verbs in
(9a, b) have the following structures in the course of derivation:

(51) a. Johni seems [TP $t_{i}$ [T for to] be better]
    b. *it seems [CP for [TP John to be better]]

(51b) is ruled out as a violation of (31a) since (31a) requires that the
verb select a (30a)-type of TP to-IC headed by for to. Similar remarks are applicable to (33c), repeated here as (52):

(52) *It seems evidently \([CP \text{for John to be better}]\).

In (51a) Agree holds between matrix T and John in situ, valuing and deleting the \(\phi\)-features of T and the Case feature of John, and then the matrix T raises John to matrix Spec-T, thereby deleting its EPP-feature. Notice that this raising is legitimate because there is no CP barrier. Thus, the present analysis correctly predicts that the derivation results in convergence, yielding the acceptable example (9a). It is also necessary to recognize that the analysis can do without (21a, b) as well as D/S-structures (22a, b), thereby eliminating the question of how to account for the optionality or obligatoriness of (21a, b).

Fifthly, adhering to the analysis here, the to-ICs of non-verbal predicates like adjectives in (10a–c) have the following structures in the course of derivation:

(53) a. it was stupid \([TP \text{PRO [T for to] do that}]\)
b. *it was stupid \([TP \text{them [T for to] do that}]\)
c. it was stupid \([CP \text{for [TP them to do that]}]\)

It is worth noting that since (31a, b) do not apply to non-verbal predicates by definition, these predicates allow either (30a)-type of TP to-ICs or (30b)-type of CP to-ICs as a free option. In (53a) Agree holds between infinitival T and PRO, valuing and deleting the \(\phi\)-features of T and the null Case feature of PRO, thus leading the derivation to converge. In (53b) Agree does not hold between stupid and them via ECM because stupid, being an adjective, has no Case agreeing capacity, thus causing the derivation to crash with the Case feature of them unvalued and undeleted. In (53c) Agree holding between for and them values and deletes the \(\phi\)-features of for and the Case feature of them, thus inducing the derivation to converge (cf. fn. 22). Hence, we correctly predict that (10a, c) are acceptable but (10b) is not. As a result, the analysis advanced can account for the fact that only the to-ICs of non-verbal predicates compel their lexical subjects to follow for as in (10c) by having recourse to (31a, b) and minimalist Case-theoretical considerations.

4.4. Further Consequences

Here I demonstrate that the suggested analysis based on (30a, b) and (31a, b) has advantages over Henry’s (1992) in many respects. The first obvious advantage, as pointed out above, is that the analysis suits
the spirit of minimalism better in that (i) it is conducted by means of
a strictly derivational approach without resort to D/S-structures, gov-
ernment, the PRO Theorem, etc.; (ii) it succeeds in dispensing with
Henry's three operations (i.e. for-cliticization, for-deletion, and CP-
deletion); and (iii) it enables us to circumvent the questions of why
these operations are optional or obligatory according to types of predic-
cates, how many of them should apply to each predicate type, and
when they should apply, hence eliminating a theoretically undesirable
"look-ahead" property (cf. sec. 3).

The second advantage, as noted above, is that the proposed analysis,
according to which BE has for as a part of infinitival morphology in T,
can maintain (4a, b) (= (35a, b)) in both BE and SE, and hence is more
universal; that is, the complementizer for and PRO are incompatible
and COMP-less to-ICs are all TPs. Consequently, the present analysis
leads us to expect that PRO always appears in (30a)-type TP to-ICs
headed by for to. Further, as mentioned above, the analysis is prefer-
able to Henry's and resonates with the current minimalist analyses such
as Bošković (1997), Doherty (1997), Iwakura (2000), etc., in that
COMP-less to-ICs have fewer projections (i.e. TPs) than to-ICs with a
COMP (i.e. CPs). In this regard, the analysis is consistent with Boš-
ković's MSP in ruling out the superfluous CP projections.

Thirdly, the suggested analysis can give a ready account of the fact
that the negative element not is placed after the infinitival marker to in
BE negative to-ICs, as illustrated in (12a, b)/(25a, b), repeated here as
(54a, b):

(54) a. I would prefer [them for to not go].
b. *I would prefer [them for not to go].

(31b) requires the matrix verbs in (54a, b) to select (30a)-type of TP to-
ICs headed by the complex infinitival morphology for to since there is
no element that is closer to v/V than the to-ICs (cf. (49c)). Thus, we
correctly predict that (54b) is ruled out because for to, being a single
unit, cannot allow any element to intervene between for and to.26
Similar remarks are applicable to the contrast between (27a, b), repro-

---

26 It is worth noticing that not cannot appear in front of the for to unit, as exem-
plified in (i):

(i) *I would prefer [them not for to go].
(Cf. SE: I would prefer [them not to go].)
duced here as (55a, b); (55b) is excluded since the single unit for to bars any AdvP from intervening between for and to. Thus, we rightly predict that (54a) and (55a) are acceptable but (54b) and (55b) are not:

(55) a. I want [the boys definitely for to be there].
    b. *I want [the boys for definitely to be there].

Fourthly, the presented analysis can offer a ready account of the fact that the negative element not may be positioned before or after to only when for precedes lexical to-infinitival subjects, as shown in (13a, b), duplicated here as (56a, b):

(56) a. I would prefer very much [for them not to go].
    b. I would prefer very much [for them to not go].

In (56a, b) the matrix verbs select (30b)-type of CP to-ICs with for-affixation onto C. According to (31b), this is not problematic, as pointed out in (50a, b). Since for in (56a, b), unlike in (54a, b) and (55a, b), does not constitute part of a complex infinitival morphological unit with to, it follows that not is allowed to intervene between for and to, as in (56a). Accordingly, we can correctly predict that (56a) as well as (56b) is acceptable.

Fifthly, the proposed analysis can readily rule out examples like (28a, b), repeated here as (57a, b), on the basis of which Henry claims that for of for to is a complementizer:

(57) a. *I want very much [for him for to get accepted].
    b. *[For him for to pay the mortgage] would be just as expensive.

Under the analysis here, for in BE to-ICs is assumed to merge in T as a part of infinitival morphology and raise to C in CP to-ICs. This en-

Pollock (1989: 375) assumes that to-infinitives have the structure [IP NP to ([Neg not]) [VP ... V ... ]] and claims that the to not versus not to ordering depends on whether to undergoes Chomsky's (1981) “rule R” according to which it adjoins to VP. Following Pollock’s claim, we can account for the deviance of examples such as (i) by saying that the for to unit in (i) (i.e. (30a)-type of TP to-ICs) does not undergo rule R in contrast with to in SE to-ICs. I assume following Chomsky (2001a, b) that such a head-movement, relevant to ordering, falls not in narrow syntax but within the phonological component. See fn. 27.

27 Under Pollock’s (1989) analysis in fn. 26, (56a, b) imply that within the phonological component to in (30b)-type of CP to-ICs may undergo or defy Chomsky’s rule R in contrast with for to in (30a)-type of TP to-ICs (cf. fn. 26). I leave to future research the question of why (56a) is preferred to (56b) and a more detailed analysis of BE negative to-ICs.
ables us to exclude examples like (57a, b) by saying that double occurrences of for are impossible. Thus, the analysis here nullifies Henry’s claim that for of for to is a complementizer.

Lastly, the analysis advanced can also account for the fact that the wh-subject undergoes extraction from within to-ICs of want-type verbs, as in (29a), repeated here as (58a):

(58) a. Who do you want [for to help you]?
    b. whoi do you want [TP t; [T for to] t; help you]]

Adhering to (31b), (58a) has the structure (58b) in the course of derivation, where want selects a (30a)-type of TP to-IC headed by for to. Wh-movement of the subject from the embedded Spec-T to the matrix Spec-C is legitimate because for of for to, being part of the infinitival morphology in T, does not create a COMP-trace effect. Consequently, we can rightly predict that (58a) is acceptable. This leads us to expect that a wh-subject cannot undergo extraction from within the CP to-ICs of want-type verbs with for-affixation onto C such as those in (50a) and (56a, b), because such extraction would yield a COMP-trace effect. This expectation is correct, as evident from the deviance of (59a) in BE:

(59) a. *Who do you want very much [for to help you]?
    b. whoi do you want very much [CP t; for [TP t; to t; help you]]

As discussed in (50a, b), (59a) has the structure (59b) in the course of derivation in which the matrix verb selects a (30b)-type of CP to-IC. Accordingly, the proposed analysis can rule out (59a) as a COMP-trace effect, as expected.

5. Conclusion

To sum up, I have explored the distribution of to-ICs in BE on the basis of Henry’s (1992) data in order to capture the generalizations in BE and SE as to the incompatibility of the complementizer for and PRO and the categorial status of COMP-less to-ICs. I have advanced a new analysis in minimalist terms (Chomsky (2000, 2001a, b)) on the basis of the idea that in BE to-ICs for is a part of infinitival morphology in T and raises to C when C is present in the same way as V raises to v. To-ICs with for stranded in T qualify as TPs, whereas those with for raised to C serve as CPs. Then, I have proposed the BE-specific conditions on the selection of the categorial status of to-ICs: verbal
predicates other than *want*-type verbs uniformly select TP to-ICs, while *want*-type verbs do so when there is no element that is closer to v/V than to-ICs. Further, I have shown a possible way of capturing the nature of the adjacency condition in terms of intervention effects within the minimalist framework.

I have also demonstrated that the analysis offered here is preferable to Henry’s in that it can capture the above-mentioned generalizations in BE and SE and in that it can dispense with all three operations Henry assumes. I have further argued that the analysis is in accordance with current minimalist analyses (Bošković (1997), Doherty (1997), Iwakura (2000), etc.) in that COMP-less to-ICs in BE are not CPs but TPs in the same way as those in SE are, consistent with economy principles such as Bošković’s MSP.

REFERENCES


Nakao, Toshio (1972) Eigoshi II (The History of English II), Taishukan, Tokyo.

Faculty of Education
Kochi University
2–5–1 Akebono-cho, Kochi-shi
Kochi 780–8520
e-mail: fuminori@cc.kochi-u.ac.jp, gonki@hkg.odn.ne.jp